

University of Groningen

Statistics and dynamics of the perturbed universe

Lemson, Gerard

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1995

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Lemson, G. (1995). *Statistics and dynamics of the perturbed universe*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 2

Point processes and galaxy clustering I: multiple-volume counting probabilities

1 Introduction

The distribution of galaxies in the Universe is far from uniform. The main aim of statistical cosmology is to find descriptors which may quantify the structural characteristics of this distribution. The most popular of these descriptors in the astronomical community uses the so-called N-point correlation functions. These measure the clustering properties of the point process under consideration on a microscopic level. The process is described in terms of density distributions giving the probability of N points occupying different positions. This local description has also been useful for describing the dynamical properties of galaxy clustering, through the formalism of the BBGKY hierarchy (Peebles (1980) and references therein). It is however not the most general way to describe the clustering of a random point process. In the mathematical literature, point processes are described through their counting probabilities, which give the joint probabilities for non-overlapping, finite volumes to contain certain numbers of points. For a sufficiently large family of volumes, these probabilities describe any point process completely, whereas not all point processes can be described by a local probability density structure (for a comprehensive review of random point processes see e.g. Daley & Vere-Jones, 1988, DV-J88).

In an influential paper, White (1979,W79) derived a link between these two statistical approaches for describing point processes, by showing how to express the single-volume counting probabilities in terms of the correlation functions. Tying up these two methods is important for several reasons. First, for formal reasons, the description of galaxy clustering may benefit from the general knowledge about random point processes available in the mathematical literature, and should therefore be phrased in the most general formalism possible. Second, for practical purposes, the determination of the clustering parameters in terms of the correlation functions is very limited. Only the lowest order correlation functions can be determined with any accuracy, but for describing the most prominent

features in the galaxy distribution one needs much higher orders. White showed that the counting probabilities depend on the correlation functions of all orders and thus represent a description of the clustering parameters of the point process complementary to the low-order correlation functions. Characteristic aspects of the correlation functions, for instance their supposed hierarchical form, or their scaling properties at higher orders, may therefore be easier to check using the counting probabilities than from the correlation functions itself (e.g. W79, Balian & Schaeffer, 1989). Third, the dynamical properties of gravitational clustering can not easily be described within the framework of the counting probabilities. The direct relation to the correlation functions, for which this in principle *is* possible, should allow the evolution of the counting probabilities to be followed as well. Finally, as will be shown in Chapter 3, the relation between correlation functions and counting probabilities leads to constraints, which may be used to assess specific models on their merits for describing galaxy clustering. Probably these are the only point processes of real interest for cosmology, and certainly the only ones studied so far in the cosmological literature (but see Chapter 3).

None of these aspects has been fully worked out yet. In this chapter we will concentrate on the first of these by completing the relation between the counting probabilities and the correlation functions. The complete description of a point process in terms of counting probabilities requires knowledge of all the multi-volume counting probabilities. We need those for instance to be able to make the transition back from the counting probabilities to the correlation functions. This is not possible from the single-volume probabilities alone. For describing certain aspects of the galaxy distribution one also needs the correlation properties of finite volumes. For instance, the correlation function of Abell clusters is a correlation between volumes containing greater than average amounts of galaxies. It should be possible to understand the high amplitude of these cluster correlation functions (Bahcall & Soneira, 1983) from the underlying point process alone, without needing the description of the early stages of this process as in Kaiser (1984). The practical description of clustering should also benefit from knowledge of multi-volume probability functions. Just as the three-point function provides information complementary to the two-point function, so will two- and three-volume counting probabilities allow to discriminate point processes that have corresponding single volume counting probabilities.

In this chapter, I have generalized the results obtained by White: the relation between the correlation functions and counting probabilities is extended to cover essentially any number of volumes. It is shown that the resulting multi-volume counting probabilities are generated by the void probability functions, which give the probability that the volumes are empty. It was one of the main results of W79 to show this for the single-volume counts. Using these results on the probability structure of point processes, various related quantities are calculated, which are of direct importance both for practical applications and for further theoretical investigations. Multi-volume moments are derived and later applied to lead to a new and somewhat modified derivation of the Limber equation, which relates the angular correlation function of galaxies on the sky to the spatial correlation functions. The continuum limit of point processes is studied using the so-called characteristic function, which is shown to be closely related to the void probability function. After deriving the characteristic function for overlapping volumes, and using the continuum limit for Gaussian

processes, we can show that the corresponding density probability distribution for overlapping volumes is also Gaussian. This leads to a new derivation of Bower's (1991) results on the cloud-in-cloud problem for Press-Schechter (1974) theory. The chapter concludes with a summary and a discussion of possible further uses of the framework developed here.

2 Derivation multi-volume counting probabilities

2.1 Notation

The derivation of the multi-volume counting probabilities as presented here generalizes the derivation of those for a single volume in W79. For that reason I will repeat the derivation of the single volume counting probabilities in section 2.2, using a slightly different notation which will be introduced here.

As is usual in probability theory (e.g. Grimmett & Stirzaker, 1992), a *sample space* Ω of possible outcomes or events is defined. In the present discussion, these events are the possible outcomes of sampling a point process with subvolumes of all possible sizes and shapes. This leads to events of the following forms

- X_i : there is a galaxy within the infinitesimal volume dX_i , at \mathbf{x}_i .
- $\neg X_i$: there is no galaxy within the volume dX_i at \mathbf{x}_i .
- $\Phi_N(V)$: there are exactly N galaxies in the finite volume V ,
possibly in addition to galaxies at explicitly mentioned points.

This last statement defines the single volume probability for finding N galaxies in a volume V . These statements are elements of the set of all events Ω . A useful property of Ω is the relation $\neg \mathcal{Y} = \Omega - \mathcal{Y}$, where \mathcal{Y} may be any statement. A probability structure on the set of statements/events is defined simply by the probability that the statement is true. For example, one may define the probability of the truth of composite statements such as

- $P(X_1 \wedge \neg X_2)$: probability that there is one galaxy in dX_1 at \mathbf{x}_1
and no galaxy in dX_2 at \mathbf{x}_2 .
- $P(X_1 \wedge X_2 \wedge \Phi_N(V))$: probability that there are exactly $N + 2$ points in
 $V \cup dX_1 \cup dX_2$, two of which are in the volumes
 dX_i at $\mathbf{x}_i, i = 1, 2$.

Some important properties of the resulting probability structure are :

- (i) $P(\Omega) = 1$
- (ii) $P(\Omega \wedge \mathcal{Y}) = P(\mathcal{Y})$, where \mathcal{Y} may be any statement
- (iii) $P((\mathcal{Z} - \mathcal{Y}) \wedge \mathcal{Y}_1 \wedge \mathcal{Y}_2 \wedge \dots) = P(\mathcal{Z} \wedge \mathcal{Y}_1 \wedge \dots) - P(\mathcal{Y} \wedge \mathcal{Y}_1 \dots) \geq 0$ when $\mathcal{Y} \subset \mathcal{Z}$

Formally, the probabilities $P(X_i)$ should vanish since the sampling volume dX_i is infinitesimally small. The importance of these probabilities lies in the definition of the N-point correlation functions, $\xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$ (note that the standard notation for the correlation

functions is used, with ξ_i instead of w_i as used in W79):

$$\begin{aligned}
P(X_1) &= n\xi_1(\mathbf{x}_1)dX_1 \\
P(X_1 \wedge X_2) &= n^2(\xi_1(\mathbf{x}_1)\xi_1(\mathbf{x}_2) + \xi_2(\mathbf{x}_1, \mathbf{x}_2))dX_1dX_2 \\
P(X_1 \wedge X_2 \wedge X_3) &= n^3(\xi_1(\mathbf{x}_1)\xi_1(\mathbf{x}_2)\xi_1(\mathbf{x}_3) + \xi_1(x_1)\xi_2(\mathbf{x}_2, \mathbf{x}_3) + \xi_1(x_2)\xi_2(\mathbf{x}_1, \mathbf{x}_3) \\
&\quad + \xi_1(x_3)\xi_2(\mathbf{x}_1, \mathbf{x}_2) + \xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3))dX_1dX_2dX_3 \\
&\text{etc.}
\end{aligned} \tag{2.1}$$

In these equations, n denotes the average density of the point process. The functions $\xi_i(\mathbf{x}_1, \dots, \mathbf{x}_i)$ are symmetric in all their arguments. For describing galaxy clustering, we are only interested in point processes that are both homogeneous and isotropic. For such point processes the correlation functions depend only on $|\mathbf{x}_i - \mathbf{x}_j|$. For future convenience we define $\xi_0 \equiv 0$, while from the definition it is clear that $\xi_1(\mathbf{x}_1) \equiv 1$.

The general N -point probability, $P(X_1 \wedge \dots \wedge X_N)$ is the sum of all terms of the form

$$n^N \xi_i(\mathbf{x}_{n_1}, \dots, \mathbf{x}_{n_i}) \xi_j(\mathbf{x}_{n_{i+1}}, \dots, \mathbf{x}_{n_{i+j}}) \dots \xi_k(\mathbf{x}_{n_{N-k+1}}, \dots, \mathbf{x}_{n_N}) \tag{2.2}$$

with $i + j + \dots + k = N$. In this sum, $\{\mathbf{x}_{n_1}, \dots, \mathbf{x}_{n_N}\}$ is any permutation of $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, while terms that are the same under the various symmetries appear only once. For instance, since $\xi_3(\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_1)\xi_2(\mathbf{x}_5, \mathbf{x}_3) = \xi_2(\mathbf{x}_3, \mathbf{x}_5)\xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)$, only one of these enters the expression for $P(X_1 \wedge \dots \wedge X_5)$

2.2 N=1 : single volume

To derive expressions for the multi-volume counting probabilities in section 2.3 we will use results from the derivation of the single-volume probabilities as presented in W79. In this section I will therefore follow that derivation using the notation presented in the previous section. The main steps involved are illustrated by the derivation of the void probability function, $P(\Phi_0(V))$. The volume V is divided in a great number of disjunct subvolumes, dX_1 to dX_N , which together span V . We may now write

$$P(\Phi_0(V)) = P(\neg X_1 \wedge \neg X_2 \wedge \dots \wedge \neg X_N) \text{ as } N \rightarrow \infty. \tag{2.3}$$

With the properties (i) to (iii) from the previous section, we can expand this expression for the void probability function as follows:

$$\begin{aligned}
P(\Phi_0(V)) &= P(\Omega - X_1 \wedge \dots \wedge \Omega - X_N) \\
&= P(\Omega \wedge \Omega \wedge \dots \wedge \Omega) - \sum_{1 \leq i \leq N} P(\Omega \wedge \Omega \wedge \dots \wedge X_i \wedge \dots \wedge \Omega) \\
&\quad + \sum_{1 \leq i < j \leq N} P(\Omega \wedge \Omega \wedge \dots \wedge X_i \wedge \dots \wedge X_j \wedge \dots \wedge \Omega) - \dots \\
&\quad + (-1)^N P(X_1 \wedge X_2 \wedge \dots \wedge X_N) \\
&= 1 - \sum_{1 \leq i \leq N} P(X_i) + \sum_{1 \leq i < j \leq N} P(X_i \wedge X_j) - \sum_{1 \leq i < j < k \leq N} P(X_i \wedge X_j \wedge X_k) + \\
&\quad \dots + (-1)^N P(X_1 \wedge X_2 \wedge \dots \wedge X_N).
\end{aligned} \tag{2.4}$$

The relations above are true *almost surely*, i.e. with probability 1, in the limit where the number of cells is large enough that the probability that there are two or more points in a cell vanishes. In that limit it is true that $\neg X_i = \Omega - X_i$. In the case of a point process containing point clusters for instance, this condition is not met. Examples of such processes will be encountered in Chapter 3.

In the continuum limit, $N \rightarrow \infty \wedge dX_i \rightarrow 0$, the discrete sums in Eq. 2.4 approach the integrals:

$$\begin{aligned} P(\Phi_0(V)) &= 1 - \int_V P(X_1) + \frac{1}{2!} \int_V \int_V P(X_1 \wedge X_2) \\ &\quad - \frac{1}{3!} \int_V \int_V \int_V P(X_1 \wedge X_2 \wedge X_3) + \dots \quad . \end{aligned} \quad (2.5)$$

The factors $1/n!$ are needed to correct for overcounting in the symmetric integrals compared to the asymmetric summation

$$\sum_{i < j} P(X_i \wedge X_j) \longrightarrow \frac{1}{2!} \int_V \int_V d\mathbf{x}_i d\mathbf{x}_j \text{ etc.} \quad (2.6)$$

Substituting the expansions for P in terms of correlation functions from Eq. 2.1, yields for the contribution of the N -point probability, $P(X_1 \wedge \dots \wedge X_N)$, all possible products of the form

$$\prod_{i=1,p} I_{j_i}^{k_i} \text{ with } \sum_{i=1,p} j_i k_i \equiv N \quad , \quad (2.7)$$

where we define

$$I_j \equiv \frac{n^j}{j!} \int_V \dots \int_V d\mathbf{x}_1 \dots d\mathbf{x}_j \xi_j(\mathbf{x}_1, \dots, \mathbf{x}_j) \quad . \quad (2.8)$$

The coefficient of this product in the expansion for $P(\Phi_0(V))$ is

$$\frac{(-1)^N}{N!} \frac{N!}{(i_1!)^{k_1} \dots (i_p!)^{k_p} k_1! \dots k_p!} (i_1!)^{k_1} \dots (i_p!)^{k_p} = \frac{(-1)^N}{k_1! \dots k_p!} \quad . \quad (2.9)$$

This is the same coefficient as appears in the expansion of $\exp(-I_1 + I_2 - I_3 + \dots)$, from which we conclude

$$P(\Phi_0(V)) = \exp \left(\sum_{j=1}^{\infty} \frac{(-nV)^j}{j!} \bar{\xi}_j(V) \right) \quad . \quad (2.10)$$

Here we have defined

$$\bar{\xi}_j(V) \equiv \frac{1}{V^j} \int_V \dots \int_V d\mathbf{x}_1 \dots d\mathbf{x}_j \xi_j(\mathbf{x}_1, \dots, \mathbf{x}_j) \quad . \quad (2.11)$$

This result shows, that the void probability function depends on the correlation functions of all orders. This is why one may consider this statistic as complementary to the low-order correlation functions, for describing the non-linear clustering properties of the galaxy distribution. The void probability function plays an even more fundamental role however, as will be shown next.

Analogous to the derivation of the void probability function one may derive expressions for probabilities like $P(Y_1\Phi_0(V))$ or $P(Y_1Y_2\Phi_0(V))$, which are needed for calculating $P(\Phi_1(V))$ and $P(\Phi_2(V))$. For instance, analogously to Eq. 2.5, we can derive

$$P(Y_1\Phi_0(V)) = P(Y_1) - \int_V dX_1 P(Y_1 \wedge X_1) + \frac{1}{2!} \int_V \int_V dX_1 dX_2 P(Y_1 \wedge X_1 \wedge X_2) - \dots \quad (2.12)$$

Abusing the notation somewhat, the terms dX_i are here written under the integral sign to explicitly show over which variables the integrals range. Substituting the expansions from Eq. 2.1 in this expressions leads to terms resembling those in Eq. 2.7, but with a term $k_i J_{j_i}(Y_1)$ substituted once for each factor I_{j_i} , where

$$J_i(Y_i) \equiv \frac{n^i}{1!} \int_V \dots \int_V d\mathbf{x}_2 \dots d\mathbf{x}_i \xi_i(\mathbf{y}_1, \mathbf{x}_2, \dots, \mathbf{x}_i) . \quad (2.13)$$

One term $I_{i_1}^{k_1} \dots I_{i_p}^{k_p}$ yields p such terms. A typical term in the expansion will be

$$J_{i_j} I_{i_1}^{k_1} \dots I_{i_j}^{k_j-1} \dots I_{i_p}^{k_p} . \quad (2.14)$$

The coefficient corresponding to this term changes to

$$\frac{(-1)^{N-1}}{(N-1)!} \frac{k_j i_j (N-1)!}{(i_1!)^{k_1} \dots (i_p!)^{k_p} k_1! \dots k_p!} (i_1!)^{k_1} \dots (i_j!)^{k_j-1} \dots (i_p!)^{k_p} = \frac{k_{i_j}}{(i_j-1)!} \frac{(-1)^{N-1}}{k_1! \dots k_p!} . \quad (2.15)$$

Substituting $k_i J_{j_i}$ for I_{j_i} in the product in Eq. 2.7 and changing the coefficient as indicated, is equivalent to differentiating each product $I_{i_1}^{k_1} \dots I_{i_p}^{k_p}$ by all the I_j for $j = 1 - \infty$, and multiplying by the factor $-J_i(Y_1)/(i-1)!$. Since this is true for all the individual terms in the expansion, it is true for the sum as a whole as well. As this sum was equal to $\exp(-I_1 + I_2 - I_3 + \dots)$, we find

$$\begin{aligned} P(Y_1\Phi_0(V)) &= \sum_{i=1}^{\infty} \frac{\partial \exp(-I_1 + I_2 - I_3 + \dots)}{\partial I_i} \frac{-J_i(Y_1)}{(i-1)!} \\ &= \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{(i-1)!} J_i(Y_1) P(\Phi_0(V)) \\ &= n \left[\sum_{i=1}^{\infty} \frac{(-n)^{i-1}}{(i-1)!} \int_V \dots \int_V d\mathbf{x}_2 \dots d\mathbf{x}_i \xi_i(\mathbf{y}_1, \mathbf{x}_2, \dots, \mathbf{x}_i) \right] P(\Phi_0(V)) . \end{aligned} \quad (2.16)$$

Defining

$$W_i(\mathbf{y}_1, \dots, \mathbf{y}_i; V) \equiv \sum_{j=0}^{\infty} \frac{(-n)^j}{j!} \int_V \dots \int_V d\mathbf{x}_1 \dots d\mathbf{x}_j \xi_{i+j}(\mathbf{y}_1, \dots, \mathbf{y}_i, \mathbf{x}_1, \dots, \mathbf{x}_j) , \quad (2.17)$$

one may show that more generally

$$\begin{aligned} P(Y_1 \wedge \dots \wedge Y_i | \Phi_0(V)) &\equiv \frac{P(Y_1 \wedge \dots \wedge Y_i)}{P\Phi_0(V)} \\ &= n^i \left[\sum_{\{i_1, \dots, i_p\}} \prod_{j=1}^p W_{i_j}(\cdot) \right] dY_1 \dots dY_p , \end{aligned} \quad (2.18)$$

where the term between square brackets is of the same form as the expressions defining the N -point correlation functions in eq (2.1), except that for each ξ_i the corresponding W_i is substituted (see W79).

The single volume counting probabilities can be determined from

$$P(\Phi_N(V)) = \frac{1}{N!} \int_V \dots \int_V d\mathbf{y}_1 \dots d\mathbf{y}_N P(Y_1 \wedge \dots \wedge Y_N \wedge \Phi_0(V)) . \quad (2.19)$$

One of the major results of W79 is that these single-volume counting probabilities are generated by the void probability function, which can be easily derived from the previous results. The form of this generating relation is

$$P(\Phi_N(V)) = \frac{(-n)^N}{N!} \frac{\partial^N}{\partial n^N} P(\Phi_0(V)) \quad (2.20)$$

with $P(\Phi_0(V))$ given by Eq. 2.10.

In Chapter 3 this relation will be used as a constraint on specific models for the counting probability distribution. In the following section I will derive the generalization of Eq. 2.20 to multiple volumes.

2.3 $N \geq 2$: multi-volume counting probabilities

The previous sections have prepared the ground such that we may now generalize the relation between the correlation functions and the single-volume counting probabilities, to cover essentially any number of finite volumes. The main steps in the generalization to multiple volumes can be extracted from the following derivation for $P(\Phi_N(V_1) \wedge \Phi_M(V_2))$, where V_1 and V_2 are two non-overlapping volumes. Following the methods of the previous section we determine $P(X_1 \wedge \Phi_0(V_1) \wedge Y_1 \wedge \Phi_0(V_2))$, where $X_1 \in V_1$ and $Y_1 \in V_2$. Defining $V = V_1 \cup V_2$, we get

$$\begin{aligned} P(X_1 \wedge \Phi_0(V_1) \wedge Y_1 \wedge \Phi_0(V_2)) &= P(X_1 \wedge Y_1 \wedge \Phi_0(V)) \\ &= n^2 \left[P(X_1 | \Phi_0(V)) P(Y_1 | \Phi_0(V)) \right. \\ &\quad \left. + \sum_{i=0}^{\infty} \frac{(-n)^i}{i!} \int_V \dots \int_V d\mathbf{z}_1 \dots d\mathbf{z}_i \xi_{i+2}(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1, \dots, \mathbf{z}_i) \right] P(\Phi_0(V)) . \end{aligned} \quad (2.21)$$

Clearly, the general expression for probabilities of the form $P(X_1 \wedge \dots \wedge X_N \wedge Y_1 \wedge \dots \wedge Y_M | \Phi_0(V))$ is equal to those for a single volume; the fact that in general the volume V will consist of two disconnected subvolumes does not change that. The differences appear when determining the multi-volume counting probabilities themselves. These are defined as follows,

$$P(\Phi_N(V_1) \Phi_M(V_2)) = \frac{1}{N!M!} \int_{V_1} \dots \int_{V_1} dX_1 \dots dX_N \int_{V_2} \dots \int_{V_2} dY_1 \dots dY_M P(X_1 \wedge \dots \wedge X_N \wedge Y_1 \wedge \dots \wedge Y_M \wedge \Phi_0(V)) . \quad (2.22)$$

The important generalization w.r.t. the single volume probabilities is that here the integrals over X_i and Y_i range only over the sub-volumes, V_1 and V_2 respectively, instead of over the complete volume V . This expression by itself is not very usefull, but I will show that it is possible to reduce it to one similar to the generating relation Eq. 2.20. To this end it is useful to expand the functions W_i , defined in Eq. 2.17, such that the integrations over the two subvolumes, V_1 and V_2 , are explicit. We use the following identity which is valid when $V_1 \cap V_2 = \emptyset$ and f_n is any symmetric function of n variables :

$$\underbrace{\int \dots \int}_{(V_1 \cup V_2)^n} f_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{k=0}^n \binom{n}{k} \underbrace{\int \dots \int}_{V_1^k} \underbrace{\int \dots \int}_{V_2^{n-k}} f_n(\mathbf{x}_1, \dots, \mathbf{x}_n). \quad (2.23)$$

We may then write

$$W_i(\mathbf{x}_1, \dots, \mathbf{x}_i; V_1 \cup V_2) = \sum_{j=0}^{\infty} \sum_{k=0}^j W_{i;jk}(\mathbf{x}_1, \dots, \mathbf{x}_j; V_1, V_2), \quad (2.24)$$

where

$$W_{i;jk}(x_1, \dots, dx_i; V_1, V_2) = \frac{(-n_1)^k}{k!} \frac{(-n_2)^{j-k}}{(j-k)!} \underbrace{\int_{V_1} \dots \int_{V_1}}_k dx_{i+1} \dots dx_{i+k} \underbrace{\int_{V_2} \dots \int_{V_2}}_{j-k} dx_{i+k+1} \dots dx_{i+j} \xi_{i+j}(x_1, \dots, x_{i+j}). \quad (2.25)$$

Here we have defined two variables, n_1 and n_2 , which from now on must be formally considered distinct, even though in the end they must be assigned the same value, n . For these functions we derive

$$\underbrace{\int_{V_1} \dots \int_{V_1}}_n dx_{i-n-m+1} \dots dx_{i-m} \underbrace{\int_{V_2} \dots \int_{V_2}}_m dx_{i-m+1} \dots dx_i W_{i;jk}(x_1, \dots, x_i; V_1, V_2) = (-1)^{n+m} \frac{\partial^n}{\partial n_1^n} \frac{\partial^m}{\partial n_2^m} W_{i-n-m;j+n+m,k+n}(x_1, \dots, x_{i-n-m}; V_1, V_2), \quad (2.26)$$

which reduces to

$$\underbrace{\int_{V_1} \dots \int_{V_1}}_k dx_1 \dots dx_k \underbrace{\int_{V_2} \dots \int_{V_2}}_{i-k} dx_{k+1} \dots dx_i W_i(x_1, \dots, x_i; V_1 \cup V_2) = (-1)^i \frac{\partial^k}{\partial n_1^k} \frac{\partial^{i-k}}{\partial n_2^{i-k}} W_0(V_1 \cup V_2). \quad (2.27)$$

In this last expression, W_0 should be expanded as in Eq. 2.24. One finally obtains

$$P(\Phi_N(V_1) \wedge \Phi_M(V_2)) = \frac{(-n_1)^N}{N!} \frac{(-n_2)^M}{M!} \frac{\partial^{N+M}}{\partial n_1^N \partial n_2^M} P(\Phi_0(V_1) \wedge \Phi_0(V_2)) \quad V_1 \cap V_2 = \emptyset \quad (2.28)$$

with

$$\begin{aligned}
P(\Phi_0(V_1) \wedge \Phi_0(V_2)) &= \\
&\exp \left[\sum_{i,j=0}^{\infty} \frac{(-n_1)^i}{i!} \frac{(-n_2)^j}{j!} \int_{V_1} \dots \int_{V_1} d\mathbf{x}_1 \dots d\mathbf{x}_i \int_{V_2} \dots \int_{V_2} d\mathbf{y}_1 \dots d\mathbf{y}_j \xi_{i+j}(\mathbf{x}_1, \dots, \mathbf{x}_i, \mathbf{y}_1, \dots, \mathbf{y}_j) \right] \\
&= \exp \left[\sum_{i,j=0}^{\infty} \frac{(-n_1 V_1)^i}{i!} \frac{(-n_2 V_2)^j}{j!} \bar{\xi}_{i,j}(V_1, V_2) \right] \quad (2.29)
\end{aligned}$$

Equations (2.28) and (2.29) form the direct generalization to two volumes of the single-volume results (2.20) and (2.10), where, again, it must be understood that numerically $n_1 = n_2 = n$. The second equality in Eq. 2.29 defines the averaged correlation functions $\bar{\xi}_{i,j}(V_1, V_2)$. The obvious generalization of these formulae to three and more volumes will be

$$P(\Phi_{N_1}(V_1) \wedge \dots \wedge \Phi_{N_M}(V_M)) = \frac{(-n_1)^{N_1}}{N_1!} \dots \frac{(-n_M)^{N_M}}{N_M!} \frac{\partial^{N_1+\dots+N_M}}{\partial n_1^{N_1} \dots \partial n_M^{N_M}} P(\Phi_0(V_1) \wedge \dots \wedge \Phi_0(V_M)) \quad (2.30)$$

where the M-volume void probability function is expanded analogous to the definition in Eq. 2.29.

The result we have obtained so far show that the whole probability structure of point processes is contained within the void probability functions. We have shown this here only for those point processes which may be described by the correlation functions, but in DV-J88 it is shown that this result holds generally. There it is proven that two point processes for which the void probability functions coincide on a large enough family of sub-volumes are equivalent¹. An explicit expression how to obtain the whole probability distribution from the void probability, is not given there. It is this explicit expression which makes the relation to the correlation functions as derived here so useful.

2.4 Moments, characteristic functions and the continuum limit

The moments of the counting probabilities are defined by

$$\langle N_1^{k_1}(V_1) \dots N_M^{k_M}(V_M) \rangle = \sum_{N_1, \dots, N_M=0}^{\infty} N_1^{k_1} N_2^{k_2} \dots N_M^{k_M} P(\Phi_{N_1} \wedge \dots \wedge \Phi_{N_M}(V_M)) \quad (2.31)$$

These can be calculated most efficiently by interpreting the resulting sum on the right-hand side of this definition as Taylor series of derivatives of the void probability function, as for instance

$$\begin{aligned}
&\langle N_1(V_1) \dots N_M(V_M) \rangle \\
&= \sum_{N_1, \dots, N_M=0}^{\infty} N_1 N_2 \dots N_M \frac{(-n_1)^{N_1} \dots (-n_M)^{N_M}}{N_1! \dots N_M!} \frac{\partial^{N_1+\dots+N_M}}{\partial n_1^{N_1} \dots \partial n_M^{N_M}} \exp(W_0(n_1, V_1; \dots; n_M, V_M))
\end{aligned}$$

¹Strictly speaking this is only true for simple point processes, i.e. those where point clusters do not occur with probability 1.

$$\begin{aligned}
&= (-n)^M \sum_{N_1, \dots, N_M=0}^{\infty} \frac{(-n_1)^{N_1} \dots (-n_M)^{N_M}}{N_1! \dots N_M!} \frac{\partial^{N_1 + \dots + N_M}}{\partial n_1^{N_1} \dots \partial n_M^{N_M}} \left[\frac{\partial^M}{\partial n_1 \dots \partial n_M} \exp(W_0(n_1, V_1; \dots; n_M, V_M)) \right] \\
&= (-n)^M \left[\frac{\partial^M}{\partial n_1 \dots \partial n_M} \exp(W_0(n_1, V_1; \dots; n_M, V_M)) \right] \Big|_{n_i=0, \forall i} .
\end{aligned} \tag{2.32}$$

For the two-volume correlation this gives the well known result

$$\begin{aligned}
\langle N_1(V_1) N_2(V_2) \rangle &= n^2 \left[\frac{\partial W_0(n_1, n_2)}{\partial n_1} \frac{\partial W_0(n_1, n_2)}{\partial n_2} + \frac{\partial^2 W_0(n_1, n_2)}{\partial n_1 \partial n_2} \right]_{n_1=n_2=0} \\
&= n^2 \left[V_1 V_2 + \int_{V_1} d\mathbf{x} \int_{V_2} d\mathbf{y} \xi_2(\mathbf{x}, \mathbf{y}) \right] \\
&= \langle N_1(V_1) \rangle \langle N_2(V_2) \rangle [1 + \bar{\xi}_{1,1}(V_1, V_2)] .
\end{aligned} \tag{2.33}$$

A useful tool for summarizing the probability structure of the distribution and for generating its moments is the characteristic function. The characteristic function of a discrete probability distribution $\{p_n\}_{n=0}^{\infty}$ is defined by

$$\phi(t) \equiv \sum_{n=0}^{\infty} p_n e^{itn}. \tag{2.34}$$

For t real, this series converges, which is not generally true for the so-called probability generating function, which is defined by

$$g(t) \equiv \sum_{n=0}^{\infty} p_n t^n. \tag{2.35}$$

The moments of the distribution are generated as follows:

$$\langle N^k \rangle \equiv \sum_{n=0}^{\infty} n^k p_n = (-i)^k \frac{d^k \phi(t)}{dt^k} \Big|_{t=0}. \tag{2.36}$$

The appropriate concept that generalizes the notion of a characteristic function to random point processes is the characteristic functional (see e.g. DV-J88), but we will not need that formalism in all its generality here. For fixed, non-overlapping volumes, V_1 to V_N , we define multi-dimensional characteristic functions

$$\phi(t_1, \dots, t_N) = \sum_{i_1, \dots, i_N=0}^{\infty} P(\Phi_{i_1}(V_1) \wedge \dots \wedge \Phi_{i_N}(V_N)) e^{i(t_1 i_1 + \dots + t_N i_N)}. \tag{2.37}$$

Using the expression for the N-point probabilities in terms of the void probability functions, we may easily derive

$$\phi(t_1, \dots, t_N) = P(\Phi_0(V_1) \dots \Phi_0(V_N) | n_1 \rightarrow n_1(1 - e^{it_1}), \dots, n_N \rightarrow n_N(1 - e^{it_N})) \tag{2.38}$$

$$\begin{aligned}
&= \exp \left[\sum_{i_1, \dots, i_N=0}^{\infty} \frac{(n_1(e^{it_1} - 1))^{i_1}}{i_1!} \dots \frac{(n_N(e^{it_N} - 1))^{i_N}}{i_N!} \right. \\
&\times \int_{V_1} \dots \int_{V_1} d\mathbf{x}_1^{(1)} \dots d\mathbf{x}_{i_1}^{(1)} \dots \int_{V_N} \dots \int_{V_N} d\mathbf{x}_1^{(N)} \dots d\mathbf{x}_{i_N}^{(N)} \xi_{i_1+\dots+i_N}(\mathbf{x}_1^{(1)}, \dots, \mathbf{x}_{i_N}^{(N)}) \left. \right] \\
&= \exp \left[\sum_{i_1, \dots, i_N=0}^{\infty} \frac{(n_1 V_1 (e^{it_1} - 1))^{i_1}}{i_1!} \dots \frac{(n_N V_N (e^{it_N} - 1))^{i_N}}{i_N!} \bar{\xi}_{i_1, \dots, i_N}(V_1, \dots, V_N) \right] .
\end{aligned}$$

The characteristic function therefore contains the same information as the void probability function, and consequently generates the whole set of counting probabilities. When the probability generating function exists, it has the same form as the characteristic function, except for the substitution $\exp(it_j) \rightarrow t_j$. In that form the multi-volume generating function has been derived for a specific family of point processes by Szapudi & Szalay (1993). They used a path-integral formalism to derive the probability generating function for doubly-stochastic Poisson processes (see DV-J88). Here a random density field, $\lambda(\mathbf{x})$, determines the local value of the density of a Poisson point process. The probability of finding N points in a volume V is then given by $P(\Phi_N(V)) = (\bar{\lambda}(V)V)^N \exp(-\bar{\lambda}(V)V)/N!$, where $\bar{\lambda}(V) = \int_V \lambda(\mathbf{x}) d\mathbf{x}/V$. Many types of point processes, for instance Poisson cluster processes, are equivalent to doubly-stochastic point processes (DV-J88). It would be interesting to see if this is true also for all point processes which can be described completely by the full hierarchy of correlation functions.

The characteristic function is very useful for investigating the continuum limit of the probability distribution of a discrete point set, i.e. the limit where the average number of points in the volume under consideration is very large. The characteristic function for a random variable $\nu(N)$, indexed by $N = 0, 1, 2, \dots$, is defined by

$$\phi_\nu(t) = \sum_{N=0}^{\infty} e^{it\nu(N)} p_N . \quad (2.39)$$

For a linear function, $\nu = aN + b$, one obtains

$$\phi_\nu(t) = e^{ibt} \phi_N(t/a) , \quad (2.40)$$

where $\phi_N(t)$ is the characteristic function of the original distribution. Defining the fluctuation variable $\delta = N/\bar{N} - 1$ gives the corresponding characteristic function

$$\phi_\delta(t) = \phi_N(t/\bar{N}) \exp(-it) . \quad (2.41)$$

With the characteristic function for the single volume counting probabilities as given in Eq. 2.38 we obtain

$$\phi_\delta(t) = \exp \left[-it + \sum_{j=1}^{\infty} \frac{(\bar{N}(e^{it/\bar{N}} - 1))^j}{j!} \bar{\xi}_j(V) \right] . \quad (2.42)$$

In the limit $\bar{N} \rightarrow \infty$, this function approaches

$$\phi_\delta(t) \rightarrow \exp \left[\sum_{j=2}^{\infty} \frac{(it)^j}{j!} \bar{\xi}_j(V) \right] . \quad (2.43)$$

The term linear in t , which corresponds to the ‘shot-noise’ term for discrete processes, has disappeared in the continuum limit.

An interesting special case is that where the correlation functions of order three and higher vanish. In the continuum limit the characteristic function of the random variable δ then reduces to

$$\phi_\delta(t) \rightarrow \exp \left(-\frac{t^2}{2} (1/\bar{N} + \bar{\xi}_2(V)) \right) . \quad (2.44)$$

Here I have kept the term $1/\bar{N}$ in the expression for the following reason: when calculating the discrete probability distribution for such a strictly Gaussian point process, the value of $\bar{\xi}_2(V)$ may not exceed $1/\bar{N}$; otherwise negative probabilities will occur, already in $P_1 = P_0 \bar{N} (1 - \bar{N} \bar{\xi}_2)$. This problem can be cured however, by introducing a three-point correlation function with an integrated size $\bar{\xi}_3(V) \sim \bar{\xi}_2(V)/\bar{N}$. In the limiting case $\bar{N} \rightarrow \infty$ one therefore needs only an infinitesimally small three-point function, and values for $\bar{\xi}_2(V) \sim O(1)$ can be assumed without problem.

The expression for the characteristic function in Eq. 2.44 is equal to the expression one obtains for the characteristic function as defined for continuous probability distributions. There,

$$\phi(t) \equiv \int_{-\infty}^{\infty} P(\delta) e^{it\delta} d\delta , \quad (2.45)$$

which is the continuum limit of the definition for the discrete probability distribution (2.34). For the normal distribution,

$$P(\delta) = \exp(-\delta^2/2\sigma^2) / \sqrt{2\pi\sigma^2} , \quad (2.46)$$

this results in

$$\phi(t) = \phi_\delta(t) \rightarrow \exp \left(-\frac{t^2}{2} \sigma^2 \right) , \quad (2.47)$$

from which we obtain the identification $\sigma^2 = \bar{\xi}_2(V)$. Note that there is no closed expression for the Gaussian model as a discrete probability distribution. In fact, for the reasons mentioned below Eq. 2.44, such a model does not even exist.

Analogous expressions can be derived for the continuum limit of the multi-volume counting probabilities, e.g.

$$\begin{aligned} \phi_{\delta_1, \delta_2}(t_1, t_2) &\equiv \sum_{N_1, N_2=0} e^{i(t_1(N_1/nV_1-1)+t_2(N_2/nV_2-1))} P(\Phi_{N_1}(V_1) \wedge \Phi_{N_2}(V_2)) \\ &\longrightarrow \exp \left[-\frac{t_1^2}{2} \bar{\xi}_{1,1}(V_1, V_1) - \frac{t_2^2}{2} \bar{\xi}_{1,1}(V_2, V_2) - t_1 t_2 \bar{\xi}_{1,1}(V_1, V_2) \right] . \end{aligned} \quad (2.48)$$

This expression is equal to the characteristic function for the joint Gaussian probability distribution (see e.g. Bardeen et al., 1986),

$$P(\delta_1(V_1), \delta_2(V_2)) = \frac{1}{2\pi\sqrt{\bar{\xi}_{1,1}(V_1, V_1)\bar{\xi}_{1,1}(V_2, V_2) - \bar{\xi}_{1,1}^2(V_1, V_2)}} \times \\ \exp \left[-\frac{\bar{\xi}_{1,1}(V_2, V_2)\delta_1^2(V_1) + \bar{\xi}_{1,1}(V_1, V_1)\delta_2^2(V_2) - 2\bar{\xi}_{1,1}(V_1, V_2)\delta_1(V_1)\delta_2(V_2)}{2(\bar{\xi}_{1,1}(V_1, V_1)\bar{\xi}_{1,1}(V_2, V_2) - \bar{\xi}_{1,1}^2(V_1, V_2))} \right] . \quad (2.49)$$

3 Applications

3.1 The Limber equation

The Limber equation (Limber, 1953) relates the angular correlation functions of the projected distribution of galaxies on the sky to the three dimensional spatial correlation functions (see also Peebles, 1980, § 49-§ 52). I will use the formalism derived in the previous sections to derive a modified version of this equation here. The angular two-point correlation function, $\omega_2(\theta_1, \theta_2)$, is defined by the excess probability of finding pairs of points in two separated solid angles on the sky:

$$P(\theta_1 \wedge \theta_2) = n_\theta^2(1 + \omega_2(\theta_1, \theta_2))d\theta_1 d\theta_2 . \quad (2.50)$$

Here n_θ is the surface density of galaxies projected on the sky. Due to the assumed homogeneity and isotropy, $\omega_2(\theta_1, \theta_2) = \omega_2(|\theta_1 - \theta_2|)$. To determine the correlation function in practice, one usually estimates the following moment

$$\begin{aligned} \langle N_1(\theta_1)N_2(\theta_2) \rangle &= \sum_{N_1, N_2=0}^{\infty} N_1 N_2 P(\Phi_{N_1}(\Omega_1) \wedge \Phi_{N_2}(\Omega_2)) \\ &= \langle N(\theta_1) \rangle \langle N(\theta_2) \rangle (1 + \overline{\omega}_2(\theta_1, \theta_2)) \end{aligned} \quad (2.51)$$

where $\Omega_{1,2}$ are finite solid angles on the sky, at the positions θ_1 and θ_2 . The probability $P(\Phi_{N_1}(\Omega_1) \wedge \Phi_{N_2}(\Omega_2))$ for observing N_1 and N_2 points in the solid angles Ω_1 and Ω_2 on the sky, is related to the spatial probability function for the corresponding cone-shaped, spatial volumes, which have their tips at the point of observation, and which are of infinite length. This angular term is not *equal* to the spatial probability, since in practice there will be selection criteria, which allow only a finite number of galaxies to be observed within the infinite cones. The usual selection criterium is the apparent luminosity of the galaxies. Only galaxies brighter than a certain lower limit are included in the catalogues. To take this into account in calculating $\omega_2(\theta)$ from $\xi_2(r)$, one needs to know the luminosity function of the galaxies as explained in detail in Peebles (1980, § 49). Here we will simply assume that all galaxies have the same absolute luminosity, and assuming a lower limit to the apparent luminosity is equivalent to introducing a maximum size for the cones.

In the following derivation we will introduce a modification to the usual Limber equation. We will assume that the common tip of the two cones is positioned at a point from the

point process. This corresponds to the fact that we observe the distribution of galaxies on the sky from within a galaxy. This condition should formally be taken into account when inverting the Limber equation to calculate the spatial correlation function from the angular on.

The quantity we need is

$$\langle N_1(\theta_1)N_2(\theta_2) | \text{ point at origin } O \rangle = \sum_{N_1, N_2=0}^{\infty} N_1 N_2 P(\Phi_{N_1}(V_1) \wedge \Phi_{N_2}(V_2) \wedge O) / P(O) , \quad (2.52)$$

where the volumes $V_{1,2}$ are cones of finite size, with their tip at O . Using the results in section 2.4 this moment is easily calculated

$$\begin{aligned} \langle N_1(\theta_1)N_2(\theta_2) | \text{ point at origin } O \rangle = & n^3 V_1 V_2 dV \left[1 + \frac{1}{V_1} \int_{V_1} d\mathbf{x} \xi_2(\mathbf{x}, \mathbf{o}) + \frac{1}{V_2} \int_{V_2} d\mathbf{x} \xi_2(\mathbf{x}, \mathbf{o}) \right. \\ & \left. + \frac{1}{V_1 V_2} \int_{V_1} d\mathbf{x} \int_{V_2} d\mathbf{y} \xi_2(\mathbf{x}, \mathbf{y}) + \frac{1}{V_1 V_2} \int_{V_1} d\mathbf{x} \int_{V_2} d\mathbf{y} \xi_3(\mathbf{x}, \mathbf{y}, \mathbf{o}) \right] / (ndV) \quad (2.53) \end{aligned}$$

If we would have assumed a more realistic luminosity function, the measures, such as $d\mathbf{x}$, would have to be multiplied by a selection function, $\phi(\mathbf{x})$, directly related to the luminosity function. Peebles (1980) calculates this function, which only depends on $|\mathbf{x}|$.

Eq. 2.53 contains three terms that are absent from the standard Limber equation. The second and third term within the square bracket are modifications to the average number of particles in each cone, due to the occurrence of a point at the base of the cones. The fourth term is the usual integral over the volume, which also appears in the standard derivation. The last term gives the excess chance of having a pair in the cone due to the occurrence of the point at the tip of the cones, and therefore involves the spatial three-point correlation function.

To obtain the angular correlation function from Eq. 2.53, we must normalize the moment in Eq. 2.53 by the surface density as observed from O . This density will in general be different from the average spatial density n , again because the point O is not a random point in space, but an element from the point process. The angular correlation function as observed will thus be related to the spatial correlation function by

$$\omega_2(\theta_1, \theta_2) \equiv \frac{\langle N_1(\theta_1)N_2(\theta_2) | \text{ point at origin } O \rangle}{\langle N_1(\theta_1) | \text{ point at origin } O \rangle \langle N_2(\theta_2) | \text{ point at origin } O \rangle} - 1 \quad (2.54)$$

$$= \frac{\int_{V_1} d\mathbf{x} \int_{V_2} d\mathbf{y} (\xi_2(\mathbf{x}, \mathbf{y}) - \xi_2(\mathbf{o}, \mathbf{x})\xi_2(\mathbf{o}, \mathbf{y}) + \xi_3(\mathbf{o}, \mathbf{x}, \mathbf{y}))}{\int_{V_1} d\mathbf{x} \int_{V_2} d\mathbf{y} (1 + \xi_2(\mathbf{o}, \mathbf{x}))(1 + \xi_2(\mathbf{o}, \mathbf{y}))} \quad (2.55)$$

For a general luminosity function and corresponding probability measure ϕ , and assuming that the cones have negligible angular sizes, $d\Omega_1$ and $d\Omega_2$, we can rewrite this to the form

$$\omega_2(\theta) = \frac{\int_0^\infty dx \int_0^\infty dy x^2 y^2 \phi(x) \phi(y) [\xi_2(d) - \xi_2(x)\xi_2(y) + \xi_3(x, y, d)]}{\left(\int_0^\infty dx x^2 \phi(x) [1 + \xi_2(x)] \right)^2}, \quad (2.56)$$

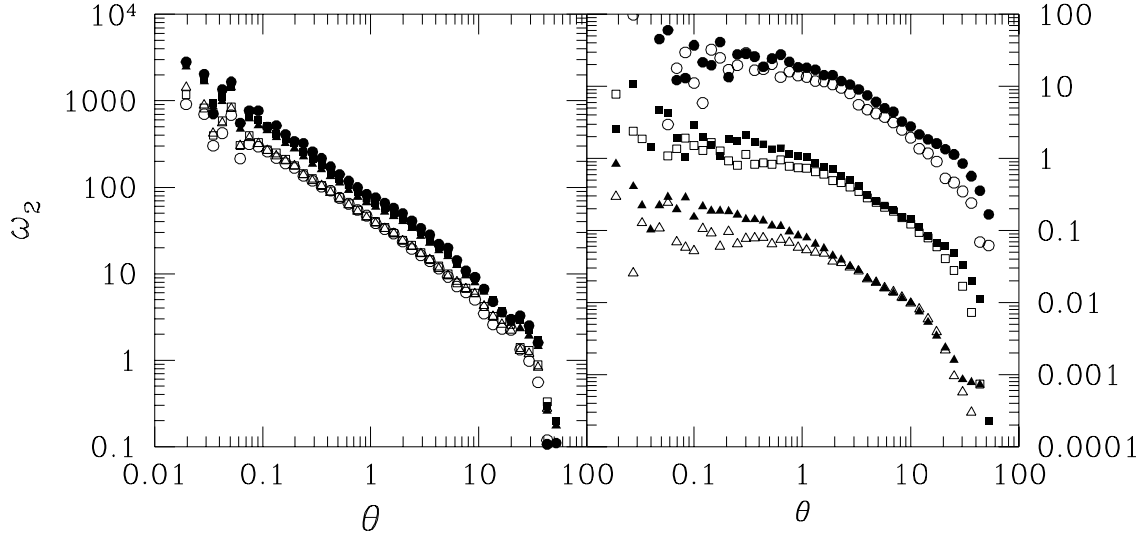


Figure 1: Two-point angular correlation functions for a Lévy-flight fractal with dimension 1.2 (left) and a Voronoi tessellation (right). The filled dots give the correlation functions determined as observed from a point of the point set itself. The open symbols give the correlation functions determined as observed from a random point. Triangles refer to the deepest sample, circles to the shallowest, as determined from the artificial apparent magnitude limit imposed on the ‘observations’. For the Voronoi tessellation, the functions for the shallow sample have been moved up by a factor 10, while those for the deep sample have been moved down by a factor 10, to separate the three different curves. On the left this has not been done, to show the scaling behaviour of the fractal point set: the amplitude of the angular correlation function is independent of depth.

where $d \equiv (x^2 + y^2 - 2xy \cos(\theta))^{1/2}$, and θ is the angular separation between θ_1 and θ_2 . One can make analytical estimates for the relative magnitudes of the correction terms. If, as is often assumed, the two-point correlation function is a power law, $\xi_2(r) = (r/r_0)^{-\gamma}$, and if the three-point function has a hierarchical form (W79; Peebles, 1980, § 54.C), $\xi_3(x, y, z) \sim \xi_2(x, y)\xi_2(x, z) + \text{cycl.}$, then upon integration over a cone of depth R_0 , the first term in the numerator will give a contribution $\sim r_0^\gamma R_0^{6-\gamma}$, while the correction terms will give a contribution of order $r_0^{2\gamma} R_0^{6-2\gamma}$. In samples which are much deeper than the correlation length, $R_0 \gg r_0$, the correction terms will vanish with respect to the standard terms. This is likewise true for the extra terms in the denominator.

To estimate the magnitude of these corrections in practical circumstances, angular correlation functions were determined for two simulated point sets, a Levy-flight fractal and a Voronoi tessellation. The construction algorithms of these point sets are described in Chapter 1 of this thesis. The points were distributed according to their respective algorithms within a sphere, and luminosities were associated to the points according to a Schechter luminosity function (Schechter, 1976). Angular correlation functions were calculated twice, once from a randomly chosen point of the sphere, not belonging to the point set, once from a randomly chosen point of the set. To investigate the dependence on the fiducial depth of the samples, the angular correlation functions were determined using variable lower limits

to the apparent luminosity. Fig. 1 shows the averaged results for five realizations of the simulated point processes, each at three different depths.

The first noteworthy feature of the fractal correlation functions is the independence of the depth of the sample. This is caused by the self-similar scaling behaviour of fractals, as explained in Peebles (1980, § 62). This scaling still appears to hold when the origin is chosen on a point of the fractal. The corresponding correlation functions have a somewhat higher amplitude than those for which the the origin was chosen at a random position. This may be explained by the fact that pairs of points have greater angular separations when viewed from close distances, which will be the case when the origin is positioned on a point of the fractal. This causes the correlation function to shift to larger angles, i.e. to the right in the diagram.

For the Voronoi tessellation, the independence of the depth of the sample does not hold. The effects of the choice of the origin are most apparent here at small angular scales, where the amplitude of the correlation function is largest when the origin is chosen on a point of the tessellation. The origin will now lie in a wall, and this will show up as a linear feature in apparent-luminosity limited maps, just as the local supercluster shows up as a linear feature in the distribution of galaxies on the sky. And just as the local supercluster is less obvious in deeper maps, so will the statistical effects of such a linear feature on the angular correlation function be diminished in deeper maps. That some trace of it remains at small scales will be due to the fact that the effective dimension of a Voronoi tessellation is 2 (see Chapter 1). The slope of the spatial correlation function is therefore -1, which leads to a flat angular correlation function when determined from a random origin. A linear feature superimposed on this flat distribution, may therefore lead to significant effects on the statistics even in deep samples.

3.2 The cloud-in-cloud problem + merging probabilities

As was seen in section 2.4, the continuum limit of the discrete counting probabilities offers an interesting approach to the description of Gaussian random fields. In this section we will apply this method to the so-called cloud-in-cloud problem of the Press-Schechter approach to non-linear clustering (Press & Schechter, 1974). In particular we will rederive the solution to this problem in an approach by Bower (1991).

In the Press-Schechter approach, the statistical properties of the initial density fluctuation field are used to infer the mass-spectrum of non-linear objects at later epochs. Here one generally uses the linear evolution of density perturbations to predict when a volume of a certain size and at a certain overdensity will reach non-linearity and recollapse to form an object at the corresponding mass scale (see Bond & Myers (1993) for an alternative approach). To determine the mass of the objects, the initial fluctuation field is smoothed at different smoothing scales, and the largest scale at which a point is still at an overdensity larger than the critical value determines the mass of the object to which it will belong.

The problem with the approach as originally developed by Press & Schechter, is that it does not explicitly take into account the possibility that structures at a smaller scale will be contained in larger structures at a later time. Bower (1991) resolved this problem by explicitly calculating the probability for such a cloud-in-cloud (see Bond et al., 1991, for an

alternative approach). Bower works within the formalism of Gaussian random field theory, and the basic step he must take is to calculate the probability that a point that has an overdensity δ_1 at smoothing scale R_1 , has an overdensity δ_2 at smoothing scale $R_2 > R_1$. Within this formalism, this corresponds to the probability $P(\delta_V, \delta_{V'})$ with $V \subset V'$, where δ_V is the overdensity in a volume $V \subset V'$. If the joint probability distribution of δ_V and $\delta_{V'}$ is Gaussian, one may easily calculate the conditional probability $P(\delta_V | \delta_{V'})$, which will also be Gaussian.

Using the formalism we have developed so far, we may calculate this cloud-in-cloud probability more directly, by calculating the joint probability of finding N points in a volume V , which is imbedded in a larger volume, V' , containing $N' \geq N$ points. We will first show that for a Gaussian point process, the joint probability $P(\Phi_N(V) \wedge \Phi_{N'}(V'))$ for overlapping volumes is also Gaussian. This follows from the characteristic function for this probability distribution:

$$\begin{aligned}
\phi_{V,V'}(t, t') &\equiv \sum_{N, N'=0}^{\infty} e^{i(tN+t'N')} P(\Phi_N(V) \wedge \Phi_{N'}(V')) \\
&= \sum_{N'=0}^{\infty} \sum_{N=0}^{N'} e^{i(tN+t'N')} P(\Phi_N(V) \wedge \Phi_{N'-N}(V' - V)) \\
&= \sum_{N=0}^{\infty} \sum_{N'=N}^{\infty} e^{i((t+t')N+t'(N'-N))} P(\Phi_N(V) \wedge \Phi_{N'-N}(V' - V)) \\
&= \sum_{N, N'=0}^{\infty} e^{i((t+t')N+t'N')} P(\Phi_N(V) \wedge \Phi_{N'}(V' - V)) \\
&= \phi_{V, V'-V}(t+t', t').
\end{aligned} \tag{2.57}$$

The characteristic function in the final line of this equation deals with non-overlapping volumes, and the expression in Eq. 2.38 can thus be used. In the continuum limit and for a Gaussian process this leads to the following characteristic function for the random variables $\delta_V \equiv N/nV - 1$

$$\begin{aligned}
\phi_{\delta_V, \delta_{V'}}(t, t') &\longrightarrow e^{-i(t+t')} \phi_{V, V'-V} \left(\frac{t}{nV} + \frac{t'}{nV'}, \frac{t'}{V'} \right) \\
&= \exp \left[-\frac{t^2}{2} \bar{\xi}_{1,1}(V, V) - tt' \bar{\xi}_{1,1}(V, V') - \frac{t'^2}{2} \bar{\xi}_{1,1}(V', V') \right]
\end{aligned} \tag{2.58}$$

This shows that the joint probability distribution of the overdensities in two volumes, one imbedded in the other, is indeed Gaussian:

$$p(\delta_V, \delta_{V'}) = \frac{1}{2\pi\sqrt{\Delta}} \exp \left[-\frac{\delta_V^2 \bar{\xi}_{1,1}(V', V') + \delta_{V'}^2 \bar{\xi}_{1,1}(V, V) - 2\delta_V \delta_{V'} \bar{\xi}_{1,1}(V, V')}{2\Delta} \right] \tag{2.59}$$

where

$$\Delta(V, V') \equiv \bar{\xi}_{1,1}(V, V) \bar{\xi}_{1,1}(V', V') - \bar{\xi}_{1,1}(V, V')^2 \tag{2.60}$$

We may now calculate the conditional probability

$$\begin{aligned} p(\delta_V | \delta_{V'}) &= p(\delta_V, \delta_{V'}) / p(\delta_{V'}) \\ &= \frac{1}{\sqrt{2\pi\sigma_{V|V'}^2}} \exp \left[-\frac{(\delta_V - \mu_{V|V'})^2}{2\sigma_{V|V'}^2} \right] \end{aligned} \quad (2.61)$$

with

$$\mu_{V|V'} \equiv \frac{\bar{\xi}_{1,1}(V, V')}{\bar{\xi}_{1,1}(V', V')} \delta_{V'} \quad (2.62)$$

and

$$\sigma_{V|V'}^2 \equiv \bar{\xi}_2(V, V) - \frac{\bar{\xi}_{1,1}^2(V, V')}{\bar{\xi}_{1,1}(V', V')} = \sigma_V^2 - \frac{\sigma_{VV'}^4}{\sigma_{V'}^2}, \quad (2.63)$$

where we have partly used Bower's notation.

To go from this result to the modified Press-Schechter theory, Bower needs to take one more important step. What one really needs is the probability that a randomly chosen volume V inside V' , is at an overdensity δ_V when the density in the larger volume is $\delta_{V'}$. Bower argues that this distribution is probably also Gaussian within a few percent, and essentially from the demand that the density for the randomly chosen volume V , given by Eq. 2.62, should equal the density of V' , he obtains

$$\bar{\xi}_{1,1}(V, V') = \bar{\xi}_{1,1}(V', V'), \quad (2.64)$$

or

$$\sigma_{V|V'}^2 = \sigma_{V'}^2. \quad (2.65)$$

In the concluding section some possible extensions of this approach are discussed.

4 Summary and conclusions

In this chapter I have developed a general theory for describing the statistical properties of stochastic point processes with the method of finite volume counting probabilities. In particular, I have related these probabilities to the hierarchy of N-point correlation functions, which forms the traditional approach for describing clustered point sets in cosmology. White (1979) has shown how the single-volume counting probabilities can be expressed in terms of all the N-point correlation functions, but for describing a stochastic point process completely one needs knowledge of the multi-volume probabilities as well. In fact, the description of galaxy clustering in terms of counting probabilities is more general than that using correlation functions. The latter can be derived from the former using the appropriate limits for the multi-volume moments.

First, expressions for the two-volume void probability function were derived in section 2.3, and there it was shown that this function generates the two-volume counting probabilities, just as single-volume probabilities are generated by the single-volume void probability function, as was first shown by White (1979). This relation can be easily generalized to more than two volumes, to show that the multi-volume void probability functions

generate all the counting probabilities, and therefore describe the point process completely. This is a well known result in the mathematical literature, but the use of correlation functions allows these generating relations to be expressed in an elegant way, which moreover is very useful for further applications. This is shown in section 2.4, where the multiple-volume moments and the characteristic function are derived. The continuum limit of the characteristic function allows us to make a connection to the statistical description of random density fields.

The expressions for the counting probabilities as derived in section 2 allow one to derive various results in a more straightforward manner. This is first shown in section 3.1, where the Limber equation is rederived and modified. Traditionally, this equation was derived without taking into account the fact that we ourselves are situated within a galaxy. This fact has consequences for the form of the Limber equation which must now be phrased as a conditional probability. This is easily achieved within the formalism developed in this chapter, and it is shown that several correction terms must be taken into account. The magnitude of these terms is estimated for two synthetic point sets, and it appears that they may not always be neglected, although, as expected, their relevance diminishes for deeper galaxy samples. Formally however, they should be taken into account when inverting the Limber equation to derive the spatial correlation functions from the angular ones.

Second, the continuum limit for a Gaussian point process is used to determine ‘cloud-in-cloud’ probabilities, for finding an overdensity δ_V in volume V situated inside volume V' , which is at an overdensity $\delta_{V'}$. Bower (1991) has used these probabilities to solve the ‘cloud-in-cloud’ problem within Press-Schechter theory (Press & Schechter, 1974). The new derivation of these probabilities is more direct than the one by Bower (1991) in that we explicitly derive the joint probabilities for overdensities in imbedded volumes, without needing the indirect way of using smoothed versions of the Gaussian random fields. In principle this approach may be generalized to more than two volumes. This would be useful for instance for the Monte-Carlo approach to the formation of dark halos as developed by Kauffmann & White (1993). These authors use the expressions for the joint probability distributions derived by Bower, to generate the merging history of virialized objects. In principle one should take into account the whole range of multi-volume probabilities, like $P(\delta_{V_3}|\delta_{V_1}, \delta_{V_2})$ etc., where $V_{2,3} \subset V_1$ and V_3 may itself be a subset of V_2 for instance. In practice these multi-volume probabilities will be more and more difficult to calculate explicitly, but treating the first few levels should be manageable. This is left to future work,

Left to future work are also two other problems which may be investigated with the formalism developed in this chapter. First, one may try to extend to multiple volumes the work by Balian & Schaeffer (1989) on the scaling model. Balian & Schaeffer made a simple assumption on the form of the correlation functions, namely that they should obey the following scaling law,

$$\xi_N(\lambda \mathbf{x}_1, \dots, \lambda \mathbf{x}_N) = \lambda^{-(N-1)\gamma} \xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N) \quad (2.66)$$

With this assumption and a few others, they were able to predict the form and scaling behaviour of the counting probabilities. Since it is very difficult to estimate the higher order correlation functions, the higher order predictions of a model such as defined in

(2.66) may be tested most effectively from the predictions for the counting probabilities. But just as the clustering properties of a point process are only partly described by the low-order correlation functions, so can the single-volume counting probabilities only give a limited constraint or confirmation for these models. It would therefore be valuable to extend this type of analysis to multi-volume counting probabilities.

A second interesting problem that may be analyzed using multi-volume counting probabilities, is to explain the cluster-cluster correlations. It appears that Abell clusters of galaxies are more strongly clustered than galaxies themselves (Bahcall & Soneira, 1983). There have been many observational developments since the original work on these cluster-cluster correlations by Bahcall & Soneira, some confirming the reality and strength of the biasing effect (Huchra et al., 1990; Peacock & West, 1992; Postman et al., 1992), others finding much smaller correlation amplitudes, or ascribing most of the effect to artificial projection effects, resulting from the subjective manner in which the Abell cluster catalogue was designed. To remedy this last situation, people have analyzed the catalogues resulting from automated plate scanning machines and objective cluster definitions (e.g. Dalton et al., Nicholl et al.). These results generally confirm the stronger correlations of clusters, albeit at a lower amplitude than originally found by Bahcall & Soneira.

Kaiser (1984) argues that one may assume that Abell clusters formed from high density areas of the primordial fluctuation field smoothed on the relevant scale. He shows that such high density regions in a Gaussian random field are more strongly clustered than the field itself, this is called natural biasing. When one assumes that the peculiar velocities of the clusters will be so small as not to have changed the distribution of the clusters by too much, the biasing of the primordial density peaks may explain the larger correlation amplitude of the present day clusters. Other authors have extended this work to the correlation functions of peaks of the initial density field (Bardeen et al., 1986) and to non-Gaussian fields.

It would clearly be desirable to understand the clustering of these regions of high galaxy density from the statistical characteristics of the galaxy distribution itself, without needing to resort to the characteristics of the primordial density perturbation field. It turns out however, to be very hard to get any firm results on this subject without the use of numerical simulations. A useful quantity to obtain is the following moment for high density regions

$$\begin{aligned} \xi_2(V_1, V_2)|_{\geq N_0} &\equiv \frac{\langle N_1(V_1)N_2(V_2) \rangle_{N_{1,2} \geq N_0}}{\langle N_1(V_1) \rangle_{N_1 \geq N_0} \langle N_2(V_2) \rangle_{N_2 \geq N_0}} - 1 \\ &\equiv \frac{\sum_{N_{1,2}=N_0}^{\infty} N_1 N_2 P(\Phi_{N_1}(V_1) \wedge \Phi_{N_2}(V_2))}{\sum_{N_1=N_0}^{\infty} N_1 P(\Phi_{N_1}(V_1)) \sum_{N_2=N_0}^{\infty} N_2 P(\Phi_{N_2}(V_2))} - 1 . \end{aligned} \quad (2.67)$$

This is the discrete analogue of the two-point correlation function for high-density regions in Gaussian random fields as calculated by Kaiser (1984). Indeed, in the continuum limit of a discrete Gaussian point process, as studied in section 3.2, Kaiser's result can be derived from the above definition. For general point processes, no useful closed expression can be extracted from Eq. 2.67, but maybe for certain models for the correlation hierarchy, like the scaling model discussed above, this will prove to be possible.

I will therefore end with a practical determination of the clustering of high density regions. In Fig. 2 I show estimates for the two-volume counting probabilities for a cosmological N-body simulation. This simulation is discussed in great detail in Chapter 4. In the

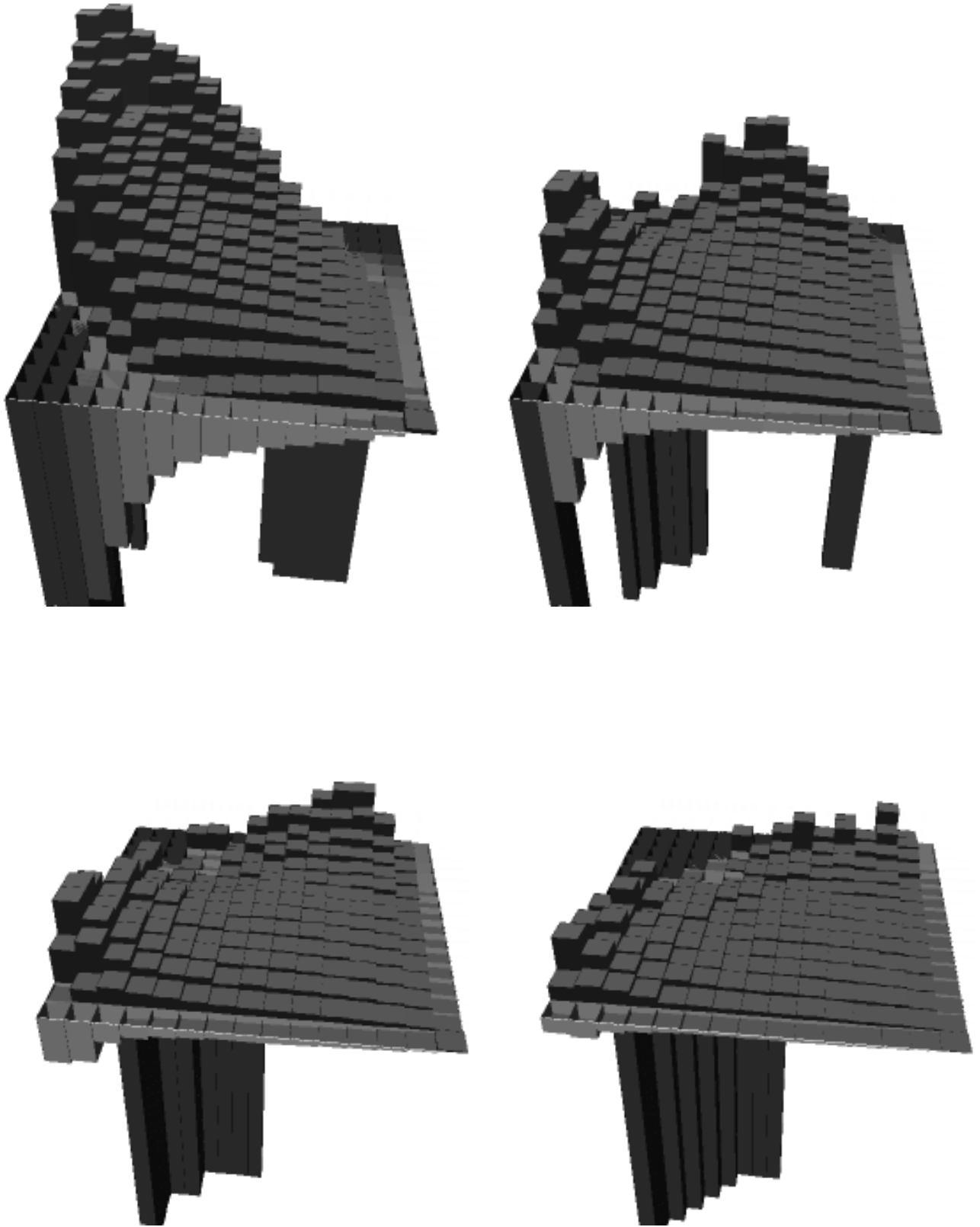


Figure 2: Estimates of the quantity $\Omega(N_1, N_2)$ defined in Eq. 2.68 for four different inter-volume separations. Shown is the logarithm of Ω as function of the counts, binned in logarithmic cells. The coordinates along the axes are, from right to left and from bottom to top: 0, 1, 2-3, 4-7, 8-15 etc. The figure on top-left is for neighbouring cells, top-right for cells with one cell in between etc.

figure 1 show estimates for the quantity

$$\Omega(N_1, N_2) \equiv \frac{P(\Phi_{N_1}(V_1) \wedge \Phi_{N_2}(V_2))}{P(\Phi_{N_1}(V_1))P(\Phi_{N_2}(V_2))} = \frac{P(\Phi_{N_1}(V_1)|\Phi_{N_2}(V_2))}{P(\Phi_{N_1}(V_1))}. \quad (2.68)$$

This quantity measures the conditional probability of finding a volume containing N_1 points, when a nearby volume contains N_2 points, relative to the unconditional probability of finding N_1 points in the volume. The estimates for this quantity are obtained by placing a cubic grid over the simulation volume and counting points in pairs of volumes. For the figures the sampling volumes were taken to be cubes 1/50 of the size of the computational box, and the four figures refer to volumes which are placed from right next to each other, to four grid-cells apart. The counts were thereafter binned logarithmically.

For neighbouring volumes we see that volumes containing larger numbers of points are more strongly correlated, while there is a strong anti-correlation between volumes containing few and volumes containing many points. In other words, when NN_2 is large, there is a large excess probability that N_1 will also be large, and this effect is stronger for larger values of N_2 . When the distance between the volumes increases, the maximum relative probability shifts to lower values of N_1 . One may compare this to the decreasing density profiles around rich clusters of galaxies. On the other hand, when N_2 is small, the probability is large that the neighbouring volume is also relatively empty, while for larger separations the maximum relative probability shifts to higher values of N_1 . A more detailed analysis of the statistical characteristics of point processes would require graphs like the one in fig. 2, for a range of volume sizes and shapes, and a larger range of separations. It would furthermore be very interesting to see whether models for the correlation functions, like the scaling model, can predict the behaviour as shown in these figures. All of this is left to future work.

References

- Bahcall, N.A., Soneira, R., 1983, ApJ, 270, 20
 Balian R., Schaeffer R., 1989, A&A, 220, 1
 Bardeen, J.M., Bond, J.R., Kaiser, N., Szalay, A.S., 1986, ApJ, 304, 15
 Bond, J.R., Cole, S., Efstathiou, G., Kaiser, N., 1991, ApJ, 379, 440
 Bond, J.R., Myers, S.T., 1993, CITA preprint, CITA/93/27
 Bower, R.G., 1991, MNRAS, 248, 332
 Daley D.J., Vere-Jones D., 1988, An Introduction to the Theory of Point Processes, Springer Verlag, New York (DV-J88)
 Dalton, G.B., Efstathiou, G., Maddox, S.J., Sutherland, W.J., 1992, ApJ, 390, L1
 Grimmett, G.R., Stirzaker D.R., 1992, Probability and Random Processes, 2nd edition, Clarendon Press, Oxford
 Huchra, J.P., Henry, J.P., Postman, M., Geller, M.J., 1990, ApJ, 365, 66
 Kaiser, N., 1984, ApJ, 284, L9
 Kauffmann, G., White, S.D.M., 1993, MNRAS, 261, 921
 Limber, D.N., 1953, ApJ, 117, 134

- Nicholl, R.C., Collins, C.A., Guzzo, L., Lumsden, S.L., 1992, MNRAS, 255, 21p
- Peacock, J.A., West, M.J., 1992, MNRAS, 259, 494
- Peebles P.J.E., 1980, The Large Scale Structure of the Universe, Princeton University Press, Princeton
- Postman, M., Huchra, J.P., Geller, M.J., 1992, ApJ, 384, 404
- Press, W.H., Schechter, P., 1974, ApJ, 187, 425
- Schechter, P., 1976, ApJ, 203, 297
- Szapudi, I., Szalay, A.S., 1993, ApJ, 408, 43
- White S.D.M., 1979, MNRAS, 186, 145 (W79)

